

Серебряная Пуля
Георапределенных Систем

Евгений Кузовлев ecommpay

Руководитель департамента эксплуатации и разработки сервисных систем ECOMMPAY IT





Кто мы такие, и почему я тут?

350+
IT-СПЕЦИАЛИСТОВ

99,99
ГАРАНТИРОВАННАЯ НАДЕЖНОСТЬ

10 млн+
ТРАНЗАКЦИЙ В ДЕНЬ

5
ДАТА-ЦЕНТРОВ ПО ВСЕМУ МИРУ

ecommpay — это про $$$



Зачем строить распределенные **системы?**

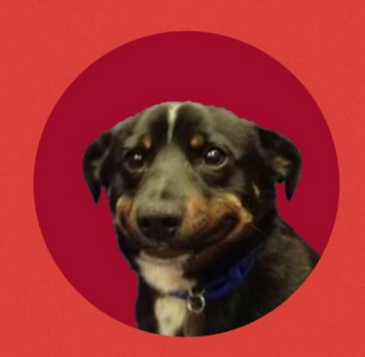



Приближать вычислительные мощности к конечному пользователю


Обеспечивать более высокий уровень надежности


Масштабировать системы и распределять нагрузку





О чем мы будем говорить?

О том, как не выстрелить себе в ногу!



Про архитектуру и синхронизацию данных



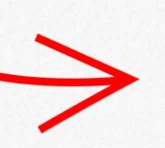
Про сеть и проблемы с сетью

О чем мы будем говорить?

О том, как не выстрелить себе в ногу!


АРХИТЕКТУРА?
КАКАЯ ЕЩЕ АРХИТЕКТУРА?





Возможные варианты синхронизации данных

Синхронная запись



Асинхронная запись



Гибридные системы





85 ms

40 ms

35 ms

10 ms

70 ms

80 ms

Сетевой лаг

World's Most Advanced Features & Un-Seen Benefits

- True Multi-master, Active-Active Cluster Read and write to any node at any time.
- Synchronous Replication No slave lag, no data is lost at node crash.
- No Master-Slave Failover Operations or Use of VIP.
- Hot Standby No downtime during failover (since there is no failover).
- Transparent to Applications Required no (or minimal changes) to the application.
- No Read and Write Splitting Needed.
- Easy to Use and Deploy

Galera cluster
пример серебряной пули

*Описание с сайта https://galeracluster.com/products/



Транснациональная корпорация. Три ЦОД: Лондон, Даллас, Сингапур.

Commit time ~0.5s.

Синхронный клиент. Latency при переключении на «резервную» ноду ~10 sec.

Перешли к асинхронной репликации через 2 месяца.

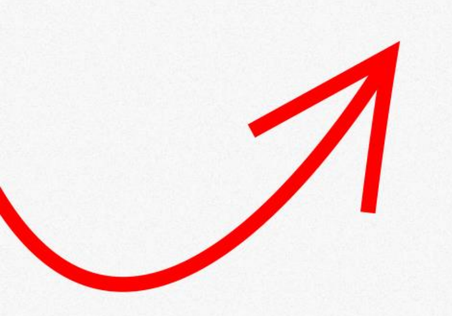


Galera cluster

пример серебряной пули



MySQL Async Master – Master





Два ЦОД: Москва и Санкт-Петербург



Асинхронная репликация



В рабочее время происходило изменение одних и тех же данных в обоих ЦОДах



Синхронная – плохо.
Асинхронная – тоже плохо.

Что делать?







Опыт из реальной жизни



Задача

построить георaспределенный биллинг


Требование 1

Максимальная скорость ответа

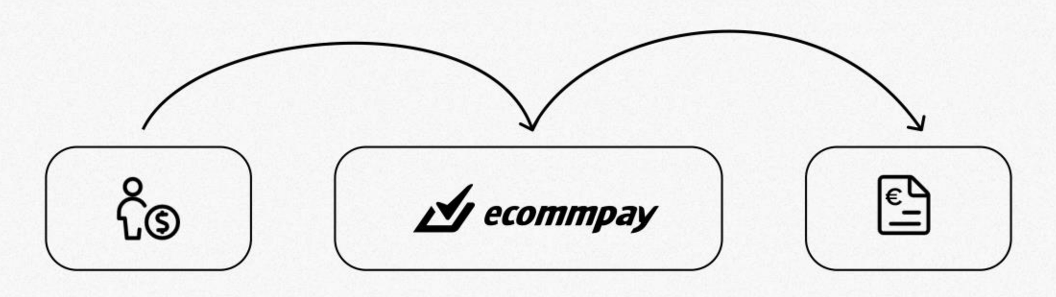

Требование 2

Синхронный контроль баланса



Типы финансовых транзакций **в биллинге**

PAY IN ТРАНЗАКЦИЯ







Типы финансовых транзакций **в биллинге**

PAY OUT ТРАНЗАКЦИЯ

€ → ecommpay → $



Опыт из реальной жизни



Задача

построить георaспределенный биллинг



Максимальная скорость ответа



Это важно только для pay-in-транзакций.



Синхронный контроль баланса



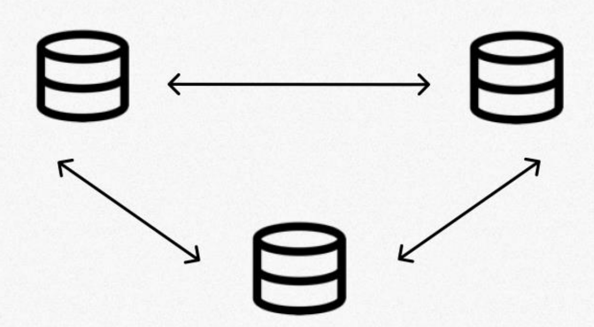
Нужен контроль недопустимости отрицательного баланса.



Архитектура системы

Balance1

Bal1_node1	**W**
Bal1_node2	R
Bal1_node3	R



Balance1

Bal1_node1	R
Bal1_node2	**W**
Bal1_node3	R

Balance1

Bal1_node1	R
Bal1_node2	R
Bal1_node3	**W**



Архитектура системы.
В итоге:

pay-in-транзакции –
максимальная скорость всегда

pay-out-транзакции –
максимальная скорость чаще всего

- Pay-in пополняют баланс, поэтому нам достаточно ответа одного ЦОД.
- Синхронное обращение к соседним нодам идет только в случае нехватки баланса.



Стройте модели синхронизации исходя из архитектуры системы.

Изменяйте архитектуру на ту, которая позволит масштабироваться на несколько ЦОД.



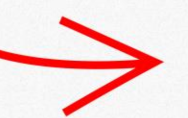
Стройте модели синхронизации исходя из архитектуры системы. **Изменяйте архитектуру на ту, которая позволит масштабироваться на несколько ЦОД.**

Но является ли синхронизация данных единственной проблемой?




РОУТИНГ





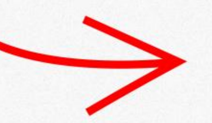
Варианты роутинга

Распределение на основании функции
(равномерное, по весам и т. д.)

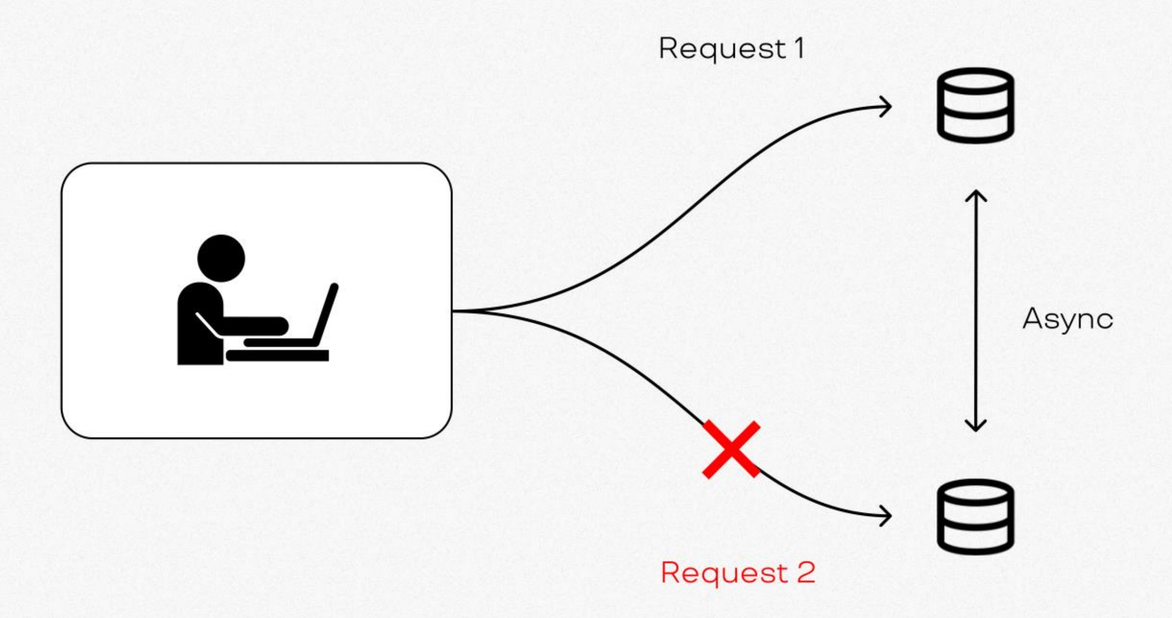
Распределение на основе параметров ЦОДов
(нагрузка, количество коннектов)

Распределение на основе параметров клиента
(IP или любые другие параметры)

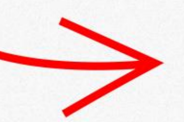




Проблематика







Варианты роутинга

Распределение на основании функции
(равномерное, по весам и т. д.)

Распределение на основе параметров ЦОДов
(нагрузка, количество коннектов)

Распределение на основе параметров клиента
(IP или любые другие параметры)

С привязкой пользователей в рамках сессии и без





Проблемы **привязки сессий**

Вследствие внутренних параметров в запросе клиента

должен присутствовать один и тот же параметр в каждом запросе.



Вследствие основных параметров клиента

перестроение всей карты при изменении конфига и перекос нагрузки.



Вследствие cookie

некуда выставлять для API-сервисов.



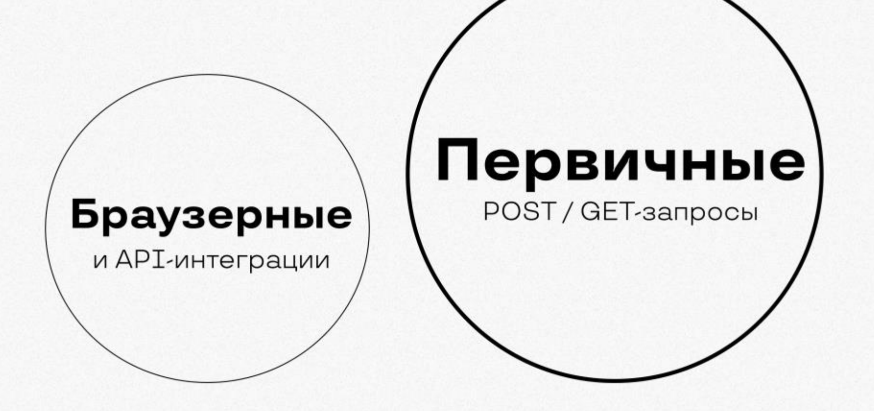
Редиректом на прямой адрес конкретной площадки

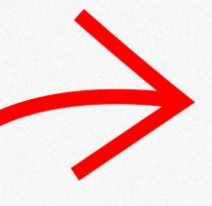
проблемы в случае первичного POST-запроса и для API-клиентов.





Наши точки входа.
Пример





Решение

комплексный роутинг на основании идентификатора проекта и идентификатора платежа на клиентской стороне.

ЦОД = f (project_id, payment_id)



Роутинг: выбор ближайшего ЦОД

AnyCast



Роутинг: выбор ближайшего ЦОД

AnyCast → **BGP**



Роутинг: выбор ближайшего ЦОД

AnyCast → **BGP** → **Собственная AS**



Собственная **AS**

IPV4-АДРЕСА ЗАКОНЧИЛИСЬ

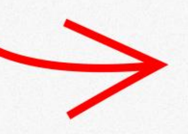
! НУЖЕН РЕЗЕРВ ПРОВАЙДЕРА

СЕТЕВИКОВ НЕТ

! РКН ТРЕБУЕТ РЕГИСТРАЦИИ

ПРОВАЙДЕР ОПЯТЬ НАКОСЯЧИЛ С FLOWSPEC





Собственная **AS и РКН**

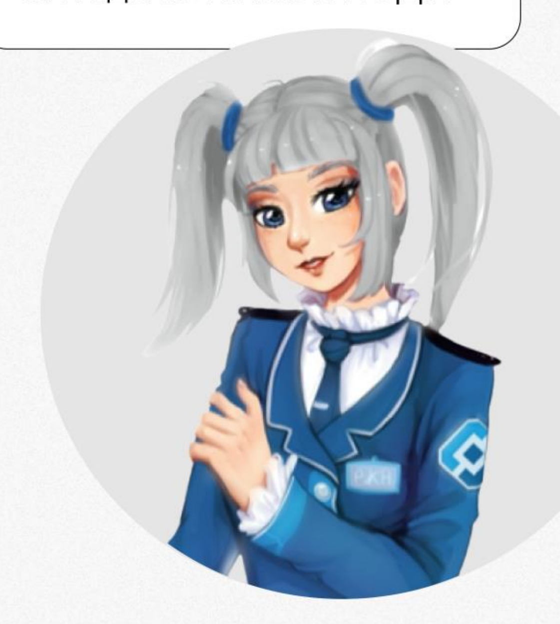
1. Передача маршрутной информации

2. Передача информации о сетевой инфраструктуре

3. Передача информации в реестр адресно-номерных ресурсов (РАНР)

4. Передача информации об инцидентах

5. Подключение к НСДИ







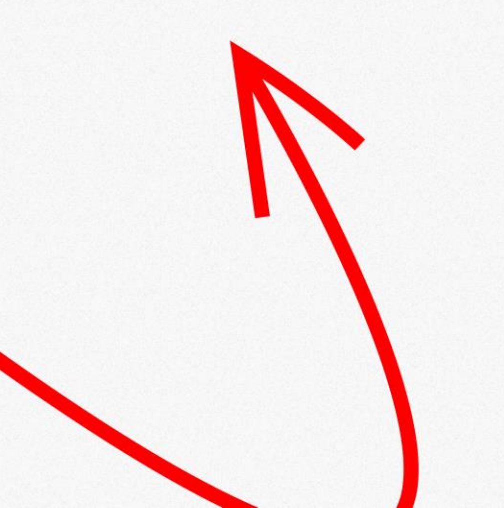
Выводы


Тип синхронизации данных и организация процесса зависит от системы и бизнес-кейсов


Не ищите готовых «серебряных пуль»


Важен вопрос роутинга пользователей для stateful-системы


Собственные пулы IP-адресов — это иной уровень контроля над архитектурой и иной уровень проблем





Спасибо !



Евгений Кузовлев

- eugene.kuzovlev@gmail.com
- e.kuzovlev@it.ecommpay.com
- +7 (905) 798-60-63
- @rafik_jan